MPRA
Munich Personal RePEc Archive

Social Interaction, Observational Learning, and Privacy: the "Do Not Call" Registry

Goh Khim Yong and Hui Kai-Lung and Png I.P.L.

April 2008

Social Interaction, Observational Learning, and Privacy: the "Do Not Call" Registry

Khim-Yong GOH, Kai-Lung HUI, and I.P.L. PNG*

April 2008

Abstract

Many empirical studies have inferred contagion in behavior from a correlation between individual behavior and the behavior of others in the same social group, rather than from any direct evidence. The correlation has been variously attributed to social interaction, word of mouth communication, and observational learning. As Manski (1993) famously observed, such correlation might be explained by peer group influence, but also, similar responses to common environmental changes. More generally, correlation in behavior raises two questions – how information is transmitted and why individuals follow the choices of others.

We address these questions in the context of subscriptions to the U.S. "do not call" registry in June-August 2003. Using a rich set of data culled from multiple sources, including longitudinal observations of household choice, we are able to separately identify

- Methods by which information is transmitted – social interaction and news media;
- Reasons why households follow the choices of others – observational learning and telemarketing diversion, and the impact of household heterogeneity on such learning and diversion.

Among methods of information transmission, social interaction was relatively more important than news media. Among reasons for contagion, telemarketing diversion was relatively more important than observational learning, while the extent of learning decreased with social heterogeneity.

* Goh and Png, National University of Singapore; Hui, City University of Hong Kong; Corresponding author: Ivan Png, School of Business, National University of Singapore, 1 Business Link, Singapore 117592, Singapore 117543.

1. Introduction

With very limited exceptions, U.S. federal law prohibits unsolicited telemarketing calls to telephone numbers on the "do not call" registry. The U.S. Federal Trade Commission (FTC) opened the "do not call" registry on June 27, 2003. Within 24 hours, over 10 million telephone numbers were registered. By October 2007, more than 145 million telephone numbers had been registered.

Relative to the U.S. population of 300 million, the explosive registration rate was phenomenal. It raises an important question of public policy: Should the government infer that the vast majority of Americans value individual privacy very highly? If so, then, the government should do more to protect privacy.

Previous research, however, has shown that people might not engage in thoughtful consideration before making choices that affect their well-being. In particular, the choices of other *nearby* individuals seem to matter – in enrolment of employees into tax-saving retirement plans (Duflo and Saez 2003), choice of health plans (Sorensen 2006), and receipt of public assistance, unemployment, and participation in public programs (Bertrand et al. 2000; Topa 2001; Conley and Topa 2002; Aizer and Currie 2004).[1]

So, could the explosive growth of "do not call" registration merely be the outcome of wide discussion about the registry – one person telling another, prompting the others to register? Or could it be due to learning by observing the choices of others rather than from direct social interaction or communication? To address these questions, we must clarify how and why individual choices affect each other.

In many empirical studies, contagion in behavior has been *inferred* from a positive statistical correlation between individual behavior and the behavior of others in the same social group, rather than from any direct evidence. The correlation has been variously attributed to social interaction, word of mouth communication, and observational learning (Glaeser et al. 1996; Dekimpe et al. 1998; Bertrand et al. 2000; Kelly and Grada 2000; Topa 2001; Talukdar et al. 2002; Topa and Conley 2002; Duflo and Saez 2003; Aizer and Currie 2004; Sorensen 2006; Beck 2007; and Moul 2007).

However, dynamic marketing efforts coupled with heterogeneity in consumer responses to marketing effort could also generate correlation in behavior (Van den Bulte and Lilien 2001; Van den Bulte and Stremersch 2004). Even if the correlation is due to individual learning from the choices of others, it is still instructive to understand the actual

[1] In the context of information privacy, Hann et al. (2007) and Hui et al. (2007) showed that people are willing to trade privacy protection for small monetary incentives.

processes – one can learn about the choices of others without any social or word of mouth influence. Alone on his island, Robinson Crusoe could choose his books, music, and movies by the *New York Times* bestseller list, *Billboard*'s charts, and weekly box office results.

Whether households subscribed to the "do not call" registry because they evaluated and considered it beneficial, because they heard from relatives and friends that it was good to register, because they inferred from the registration of others that "do not call" was a good thing, or because they read or heard about the benefits of the registry from newspapers, the Internet, or TV have distinct marketing and policy implications.

If people get their information from others, then promotion and communication (subscribe to the "do not call" registry, have more babies, be socially responsible, buy the iPhone, join Second Life) should target opinion leaders, who would then spread the message and exert influence through their social network. If people infer the value of the item from the choices of others, then these choices should be publicized (by communicating that so many millions of people have already registered / bought / joined). However, if people exhibit correlated behavior simply because they share common personal characteristics (Manski 1993; Soetevent 2006), then there is no business or policy reason to target opinion leaders or publicize the choices of others. Such efforts would have no effect.

And, finally, if people get information through newspapers, TV, the Internet, or other mass media, then communication should focus on the benefits of the product and services (avoid telemarketing calls, social responsibility is better for all, Second Life is cool) and be placed through such media.

In this study, we investigated household subscription to the "do not call" registry. Using a rich data-set compiled from multiple sources, we estimated a discrete choice model that decomposed household utility into intrinsic preferences and social factors. The data sources included the FTC, Audit Bureau of Circulation (ABC), Direct Marketing Association (DMA), Google News, U.S. Census, and the Social Capital Community Benchmark Survey (administered by the Saguaro Seminar, Harvard University). These data allowed us to identify four social factors in household registration – social interaction, communication through mass media, observational learning, and telemarketing diversion – as well as the impact of economic and social heterogeneity on such learning and diversion.

Our modeling approach and the combination of datasets, in particular, the longitudinal data on household registrations and the data pertaining to different social factors, allowed us to specifically address two questions: *how information on the "do not call" registry was* *communicated* (whether through social interaction or impersonally through mass media), and *why households followed the registrations of others* (whether due to observational learning or telemarketing diversion). To our knowledge, this is the first attempt to separate the effects of different means of communicating information, and the different motivations underlying contagion in observed behavior in a single empirical study.

We found that, among methods of information transmission, social interaction was relatively more important than news media, while among reasons for contagion, telemarketing diversion was relatively more important than observational learning, and the extent of learning decreased with social heterogeneity. People indeed had strong demand for privacy from telemarketing, and such need was more sensitive to telemarketing diversion than social factors.

The remainder of this paper is organized as follows. Section 2 reviews previous research into contagion in individual behavior. Section 3 provides a synopsis of the "do not call" registry. Section 4 introduces our research model and the various datasets and variables used. Then, we present the estimation results in Section 5, and discuss robustness in Section 6. Finally, Section 7 concludes the paper.

2. Literature Review

Individual actions and choices have been found to be contagious in many contexts. For instance, using city-level data, Glaeser et al. (1996) found high dispersion of crime rates across cities, which could not be well explained by individual or urban characteristics. They attributed the variance of crime rates across cities broadly to social interaction, but they did not characterize the specific processes of interaction.

Similarly, using U.S. census tract data, Topa (2001) and Conley and Topa (2002) showed that unemployment rates were correlated among individuals in closer social networks, as measured by education, racial and ethnic composition variables, and neighborhood boundaries. They assumed that these correlated unemployment rates arose because people shared information about job opportunities. However, correlation in unemployment rates could also be explained by common neighborhood characteristics (e.g., localized employment opportunities) or impersonal communication (e.g., some local newspapers list more job openings than others).

Aizer and Currie (2004) found that the use of public prenatal and delivery services was correlated within groups defined by race / ethnicity and neighborhoods, but they did not find evidence of information sharing. Sorensen (2006) observed that employees' choices of health plans were clustered within departments, and the department size and employees' demographic distance could explain these correlated choices. Neither study provided tangible evidence of social interaction, but rather, they inferred the "social interaction" from correlations and clustering in choices.

The same limitation applies to studies of contagious behavior in many other contexts, including education (Sacerdote 2001), purchase of consumer durables (Dekimpe et al. 1998; Talukdar et al. 2002; Goolsbee and Klenow 2003), and choices of books and movies (Beck 2007; Moul 2007; Santugini 2007). In all of these studies, correlations between individual behavior and aggregate past behavior, or behavior of others in a close social group were attributed to either social interaction, word of mouth, or observational learning, without any explicit characterization or measures of the relevant processes.

As Manski (1993) famously pointed out, using aggregate behavior to explain individual behavior is subject to the *reflection* problem. Individual choices may be correlated because they share common personal characteristics (*correlated effects*), or because individuals respond to the exogenous characteristics of others (*contextual effects*), or because individuals are influenced by the choices of others (*endogenous effects*). Any aggregate behavior may "reflect" all of these three effects, and using it as an explanatory variable may cause an upward bias in estimating the endogenous effects, i.e., contagion.

To address the reflection problem arising from analysis of aggregate behavior, researchers have applied instrumental variable estimation, specific modeling assumptions, or carefully constructed panels or experiments (see, e.g., Duflo and Saez 2003; Sorensen 2006; Moul 2007; and also, the survey by Soetevent 2006). However, to draw sharper policy implications, it would be best if the social factors that generated these correlated behaviors were directly observable.

"Observational learning" is to learn about some uncertain information from observing the choices or decisions of others without knowledge about the others persons' individual characteristics (Bikhchandani et al. 1991). In the context of waiting lists for kidney transplants, Zhang (2006) showed that, if a kidney had been rejected by patients high on the waiting list, patients further down the waiting list would draw adverse inferences from the rejections and were then more likely to reject the kidney as well.

An issue in the study of observational learning is to distinguish learning from pure information saliency, which is the pure effect of information about the item without information about the choices of others. In a nicely designed field experiment at a Beijing restaurant, Cai et al. (2007) found that when customers were informed of the five most popular dishes, demand for those items rose by 13-20%. In contrast, just informing customers of the five dishes without any ranking information had no effect on demand. However, it is not easy to construct such controlled experiments for government programs or new mass-marketed products, the adoptions of which are mostly subject to a mix of social factors including social interaction and observational learning, and also, pure information saliency.

Bertrand et al. (2000) measured social networks by the density of people in the same language group (what they called "contact availability"), and used it to weight the effect of others' receipt of public assistance on any one person's decision to receive public assistance. Their study provided rich insight into the social dynamics of welfare participation, primarily because they employed an exogenous piece of information – the number of other people that a person could speak to – as a proxy measure for inter-personal communication.

Our research generally followed Bertrand et al.'s approach, but we used more direct measures of social interaction, such as how often people visit relatives, friends, and co-workers rather than indirect measures such as language spoken. Further, we used longitudinal data to reduce the potential bias caused by using contemporaneous registration data of others as an independent variable in the estimation. Hence, we were able to minimize the reflection problem and address more precisely the questions of *how* information is communicated, and *why* the choices of others affect individual behavior.

3. "Do Not Call" Registry

The FTC opened the "do not call" registry on June 27, 2003. For the first 10 days, residents of states west of the Mississippi (including Minnesota and Louisiana) could register through the Internet and a toll-free telephone number. Residents of all other states could only register through the Internet. However, from July 7 onward, everyone could register through the Internet and telephone (FTC 2003a).

The FTC stipulated that all registrations prior to September 1, 2003 (10 weeks after the registry was opened) would be effective from October 1, 2003, while all subsequent registrations would be effective only after a 90-day waiting period (FTC 2003b). From January 1, 2005, the waiting period was cut to 31 days. Listings in the "do not call" registry were effective for five years.[2]

---

[2] At the time of writing, the FTC pledged not to drop any telephone number on the registry, even after the five-year expiration period. See http://www.ftc.gov/opa/2007/10/dnctestimony.shtm [Accessed February 20, 2008].

The registry applied to both inter-state and intra-state telemarketing calls to *residential* numbers. Any telemarketer who called a number on the registry could be fined up to $11,000. The registry did not apply to political campaigning, survey research, nonprofit and charitable organizations, and organizations with a recent commercial relationship with the consumer.

Perhaps not surprisingly, the telemarketing industry bitterly fought the federal "do not call" registry in U.S. courts.[3] On September 23, 2003, U.S. District Judge Lee R. West of Oklahoma held that the FTC had no authority to operate the "do not call" registry. Two days later, U.S. District Judge Edward W. Nottingham of Colorado held that the registry violated the constitutional right to free speech. The FTC suspended the "do not call" registry from October 1 to October 7, 2003, when the U.S. Court of Appeals for the 10th Circuit suspended the District Court orders.

On February 17, 2004, the Court of Appeals overruled the District Courts and held that the "do not call" registry was constitutional as it "targets speech that invades the privacy of the home, a personal sanctuary that enjoys a unique status in our constitutional jurisprudence."[4] Finally, on October 4, 2004, the U.S. Supreme Court declined to hear the telemarketers' appeal, which ended their legal challenge.

Prior to the federal "do not call" registry, 27 states had established state-level "do not call" registrations. Some states charged a fee for these registrations. Subsequently, 17 states merged their registries with the federal registry, while others maintained their registries in parallel with the federal registry (Varian et al. 2004).

The FTC provided us with data on all registrations, including the redacted telephone number with area code and exchange prefix (e.g., (617) 363-xxxx), and the date of registration, with the "do not call" registry between June 26, 2003 and January 6, 2006.[5]

It is useful to examine the time profile of household registrations. Figure 1(a) reports weekly registrations of fixed-line telephone numbers from Illinois. Registrations rose sharply to peak at about 700,000 in week 2, and then diminished, but peaked again in week 10, just before September 1, 2003 (recall that registrations prior to that day were effective on October

---

[3] The following review of legal actions against the "do not call" registry is based on a chronology compiled by the Electronic Privacy Information Center (http://epic.org/privacy/telemarketing/dnc/; [Accessed February 20, 2008]).

[4] "Supreme Court Upholds Do-Not-Call Registry," *Washington Post*, October 5, 2004.

[5] Varian et al. (2005) analyzed the demographics of the federal "do not call" registry at the county level, and found that registration was positively associated with household income and negatively associated with education and the presence of teenagers in the household. Varian et al. did not consider the social dynamics of "do not call" registrations.

1, 2003, while all subsequent registrations were effective only after 90 days). Registrations then tailed off until early 2005.

[Insert Figure 1 about here]

Figure 1(b) reports weekly registrations from Massachusetts, which had a somewhat different registration profile from that of Illinois. In Massachusetts, registrations also peaked at 200,000 in week 2, then diminished, but then peaked again at over 1,000,000 in week 8. The peak in week 8 corresponds to Massachusetts merging its state-level "do not call" list with the federal registry (Varian et al. 2004, Table 22). Thereafter, the registration profile of Massachusetts was similar to that of Illinois. Figure 1(c) reports weekly registrations from Texas, which had a similar registration profile as Illinois.

The above review of the chronology of the federal "do not call" registry, and the observed registration profiles in the various states (Figure 1), provide some useful guidance on sample selection. Specifically:

- We excluded telephone numbers which belonged to the federal or state government emergency, non-emergency, and directory information services. We also excluded mobile telephone numbers, as it was not possible to associate mobile phone numbers with geographical units of analysis.[6]
- We excluded registrations from any state that offered a state-level "do not call" registry. State registries were more costly and less well publicized than the federal registry (some states charged a fee, while the federal registry is free of charge). Accordingly, the households that registered on state lists would probably be those who valued privacy more. Further, among states, such as Massachusetts, that merged their lists with the federal registry, the state-registered telephone numbers were recorded in the federal registry in one batch, and were not separately identifiable. This would distort the time profile of individual household registrations. To avoid measurement errors of this nature, we focused on the 28 states that did not provide state-level "do not call" registries.[7]

[6] A tangential issue is whether the "do not call" registry provides greater benefit to fixed-line or mobile numbers. In the United States, the receiving party must pay to receive an incoming mobile call but not an incoming fixed-line call, so receiving telemarketing on a mobile number is more costly. On the other hand, the Telephone Consumer Protection Act of 1991 prohibits the use of auto dialers to call mobile numbers or leave any prerecorded message, hence telemarketers are less likely to call mobile as compared with fixed-line numbers (Varian et al. 2004).

[7] We treated the District of Columbia as a "state". A series of Chow tests rejected the null hypothesis that there was *no* structural difference in the processes leading to "do not call" registrations between

- Finally, as noted above, registrations before and after September 1, 2003 had different effective starting dates, and the “do not call” registry was subject to legal challenge between late 2003 and early 2004. To investigate the demand for the “do not call” registry and its relation to social interaction, without contamination by external events, we focused on the first nine weeks of registration.

[Insert Figure 2 about here]

Figure 2(a) reports the spatial autocorrelations (measured by the Moran’s *I* statistic) of weekly “do not call” registrations between all counties and telephone exchanges in New Jersey.[8] New Jersey comprised 21 counties and 3,003 telephone exchanges, and was the most densely populated U.S. state. At the telephone exchange level, the average spatial autocorrelation in New Jersey was 0.14 (all $p < 0.01$), whereas at the county level, the average was only 0.02, and most of the autocorrelations were insignificant.

The spatial autocorrelation was significantly stronger among telephone exchanges than among counties in New Jersey ($t = 10.66$, $p < 0.01$). Indeed, this was also the case for 19 out of 24 states with no state-level registry. This suggests that the contagion in “do not call” registrations was stronger among telephone exchanges than counties. Typically, a telephone exchange is a smaller geographical unit than a county, so, this result provides preliminary support to direct social interaction – individuals residing in closer proximity would interact relatively more with each other. This effect is likely to be weaker at the county level.

By way of contrast, Figure 2(b) reports the spatial autocorrelations for New Hampshire, which comprised 10 counties but only 514 telephone exchanges, and was relatively thinly populated compared with New Jersey. The spatial autocorrelations were mostly negative and insignificant in New Hampshire. The average spatial autocorrelations at the county (mean = –0.05) and telephone exchange (mean = –0.06) levels were not

---

the counties in states with state-level “do not call” registries and those without. Hence, it was reasonable to exclude the states with state-level “do not call” registries.

8 The Moran *I* statistic is a commonly used measure of spatial autocorrelation. It is defined as

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \cdot \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2},$$

where n is the number of counties (telephone exchanges), x_i is the number of “do not call” registrations from county (telephone exchange) i, and w_{ij} is the distance between counties (telephone exchanges) i and j, computed by the longitudes and latitudes of their centroids. Under the null hypothesis of no spatial autocorrelation, the expected value of the Moran *I* statistic is $1/(n – 1)$.

significantly different from each other ($t = 1.14$, $p = 0.26$). Figure 2(c) reports the spatial autocorrelations for Illinois, which exhibited a similar pattern as New Jersey.

The "do not call" registry is an appropriate setting to study *how* and *why* social interaction affects household choice for several reasons. The registry is maintained by FTC, and so provides the same benefit to all U.S. households. FTC does not charge any fee for registration, and so there is no price or income effect which could generate spurious observation of contagious behavior (Van den Bulte and Stremersch 2004). Also, since the registry is spearheaded by the federal government, any observed contagion in registration is not likely caused by competitive or dynamic marketing programs (unlike, for example, the medical innovations studied in Van den Bulte and Lilien (2001)).

4. Model and Data

"Do not call" registration is a discrete choice – each household simply decides whether to register its telephone number or not. Hence, we used a random utility discrete choice model to study household choice, and the preferences associated with different social factors.

Consider any household i in county j which has not yet subscribed to the "do not call" registry. Let the household's utility in period t from the "do not call" registry be

$$u_{ijt} = x'_{ijt}\beta + z'_{ijt}\alpha + \xi_{jt} + \varepsilon_{ijt}, \qquad (1)$$

where x_{ijt} is a vector of household characteristics that may affect its intrinsic preference for the registry, z_{ijt} is a vector of social factors (social interaction, communication through the mass media, observational learning, etc.), ξ_{jt} captures the unobserved "quality" of the "do not call" registry, which is common to all households in county j at time t,[9] ε_{ijt} captures household-specific random errors (e.g., idiosyncratic tastes toward the "do not call" registry, or other unobserved heterogeneity at the household level), and α and β are vectors of parameters that we estimated.

By assuming that the ε_{ijt} are independently and identically distributed with the extreme value distribution, and by normalizing the utility of households who did not subscribe to the "do not call" registry to zero, the *proportion* of households in county j who registered at time t is (McFadden 1978):

---

[9] The utility that the "do not call" registry provides to households may depend on county characteristics which are unobservable to the econometrician. For example, households in counties where retail stores offer more sales promotions or which receive more direct-mail advertising may find telemarketing calls less useful, and they may get a higher utility from subscribing to the "do not call" registry.

$$p_{jt} = \frac{e^{\delta_{jt}}}{e^0 + e^{\delta_{jt}}} = \frac{e^{\delta_{jt}}}{1 + e^{\delta_{jt}}}, \tag{2}$$

where

$$\delta_{jt} = \bar{x}'_{jt}\beta + \bar{z}'_{jt}\alpha + \xi_{jt} \tag{3}$$

is the *mean utility* (Berry 1994) obtained by households in county j at time t from subscribing to the "do not call" registry.

The proportion of households who did not register is the complement of (2),

$$1 - p_{jt} = \frac{1}{1 + e^{\delta_{jt}}}. \tag{4}$$

Taking logarithms and subtracting (4) from (2),

$$\log(p_{jt}) - \log(1 - p_{jt}) = \delta_{jt} = \bar{x}'_{jt}\beta + \bar{z}'_{jt}\alpha + \xi_{jt}. \tag{5}$$

Let M_{jt} be the entire pool of potential registrants and r_{jt} denote the actual number of registrations in county j at time t. Then, $p_{jt} = r_{jt} / M_{jt}$, and (5) simplifies to

$$\log(r_{jt}) - \log(M_{jt} - r_{jt}) = \bar{x}'_{jt}\beta + \bar{z}'_{jt}\alpha + \xi_{jt}. \tag{6}$$

By assuming that ξ_{jt} is a county-level random variable, we could estimate (6) by linear statistical procedures such as ordinary or generalized least squares.

Since "do not call" registrations were effective for five years, households need register each telephone number only once in five years. Once a number was registered, it would "exit" the market. In this regard, the "do not call" registry differed from markets for differentiated products such as automobiles, PCs, and movies (e.g., Berry et al. 1995; Hui 2004; Moul 2007), where repeat purchase does occur and must be accounted for.

Accordingly, in our model, households that have already registered would drop out from the pool of potential registrants, that is, the "market potential", M_{jt}, declines over time,[10] and hence the market potential in each period t can simply be calculated as

$$M_{jt} = M_{j0} - \sum_{k=1}^{t-1} r_{jk}.$$

Therefore, the estimation equation (6) simplifies to

$$\log(r_{jt}) - \log\left(M_{j0} - \sum_{k=1}^{t} r_{jk}\right) = \bar{x}'_{jt}\beta + \bar{z}'_{jt}\alpha + \xi_{jt}. \tag{7}$$

[10] Using data from market researcher TNS, Varian et al. (2004) calculated that, in 1999-2000, 19.7% of households had two or more fixed-line telephone numbers. Such households would bias both r_{jt} and M_{jt} upward, and so, would have little influence on the observed proportion of registrations, p_{jt}.

With the initial state of each county j, M_{j0}, we could compute all subsequent M_{jt} from the number of registrations that we observed in the following periods.[11]

Finally, we decomposed the county-level random error into four components,

$$\xi_{jt} = \gamma + u_j + d_t + v_{jt}, \tag{8}$$

where γ is an overall state-specific constant, u_j denotes time-invariant heterogeneity specific to county j, d_t captures all time-specific demand shocks which apply equally to all counties in period t, and v_{jt} captures residual errors. In our estimation, we used a random-effects specification to estimate u_j, and a set of time dummy variables to estimate d_t.[12]

4.1. Variables

Many intrinsic factors could influence a household's preference for "do not call" registration. For example, some households may want to register more than others because they have higher disposable income and so tend to receive more telemarketing calls (Varian et al. 2004), because their "cost" of registration (e.g., learning how to use the FTC's website) is lower, or because they simply have more time to read about the registry.[13]

Accordingly, we used a set of demographic variables *at the county level* to estimate households' intrinsic preferences for "do not call" registration, i.e., $\bar{x}_{jt}$ (recall from (3) and (7) that we specified our discrete choice model at the county level, and so we studied *mean* rather than individual household behavior). These county-level demographics included the average

[11] This computation of future market potential is correct only if the number of households in county j is relatively stable over time, which is likely to be the case in our study because it was limited to the first nine weeks after the "do not call" registry was opened.

[12] We had two reasons for using the random-effects specification. As described below, we relied on the Social Capital Community Benchmark Survey of the Saguaro Seminar for measures of social interaction. The Saguaro Seminar employed proportionate sampling – a variation of stratified random sampling – of U.S. population groups. This would suggest using random-effects estimation since we are ultimately interested in making unconditional or marginal inferences with respect to the entire population. The other reason was quite pragmatic – as we discuss below, several key variables, including the measures of social interaction, newspaper circulation, and telemarketing sales intensity, did not vary over time, and hence they would be "differenced out" and could not be separately identified in a fixed-effects specification. Fixed-effects estimators can be very inefficient, and lead to unreliable point estimates, when there is little "within" variance in the studied variables (Plumper and Troeger 2007), which was the case in our data. Nonetheless, we conducted some diagnostic tests of fixed-effects vs. random-effects in our estimation (see footnote 19).

[13] Using a field experiment, Hann et al. (2006) showed that spam is targeted to selected customer segments. Hence, households with different demographic characteristics may receive different amount of direct marketing solicitations. For general analyses of the strategic interactions between direct marketers and consumers, see Van Zandt (2004), Anderson and de Palma (2006), and Hann et al. (2008).

number of people in a household, median household income, percentage of people who were economically active, percentage of people above 16 years old who were employed, percentage of people below 65 years old, percentage of households with a female head, percentage of people above 25 years old with at least high school education, percentage of people who were linguistically isolated, and percentage of owner-occupied households.

Household preferences for "do not call" registration may also depend on external, social factors. Social factors can be broadly classified into "information" and "influence". When I meet my colleague at the coffee machine, she might ask, "Have you heard about the "do not call" registry?" (information), or she might say, "everybody is signing up" (information and influence), or she might say, "you should sign up" (influence). Previous research did not distinguish information from influence, rather simply lumping both into the single concept of "social interaction" (e.g., Topa 2001; Sorensen 2006).

Within the "information" category of social factors, we considered two means of communication – social interaction and news media. For social interaction, we used measures of the frequency of people interacting with neighbors, relatives, friends, and co-workers, and we multiplied these variables by the lagged cumulative *proportion* of households who had already subscribed to the "do not call" registry. These composite variables enabled us to estimate the effect of information arising from interaction with people who had already subscribed to the "do not call" registry.[14]

As for communication through news media, i.e., information saliency (Cai et al. 2007), we used the newspaper circulation in each county and the lagged number of newspaper reports of the "do not call" registry, weighted by newspaper circulation. In addition to capturing impersonal communication of information, these variables helped to control for the effect of mass marketing, which can lead to contagion-like adoption patterns (Van den Bulte and Lilien 2001). We used the lagged reports to avoid capturing the effect, in the reverse direction, of registration on news.

Within the "influence" category of social factors, we incorporated two forms of social influence in our model. One was observational learning (Bikhchandani et al. 1991 and 1998), as measured by the lagged number of newspaper articles that reported the *number* of "do not call" (DNC) registrations, weighted by newspaper circulation. To the extent that the

[14] Bertrand et al. (2000) followed a similar method to measure network effects, using the language spoken by a person as an indirect measure of social communication. Their approach did not consider that people who speak the same language might vary in participation in social activities or interaction, which is the key variable of interest. For empirical evidence, see Alesina and Ferrara (2000) and Marmaros and Sacerdote (2006).

coverage in local media – newspapers, radio, and TV – are correlated, the newspaper mentions would reflect the larger media picture as well.

Second, strategic interaction between telemarketing vendors and consumers may generate a unique type of indirect social influence – the effective exposure of households to telemarketing may change according to the number of households that have already subscribed to the "do not call" registry (Hann et al. 2008). As more households subscribe to the "do not call" registry, telemarketing calls might be diverted towards the remaining households who have not registered their numbers. This would motivate the remaining households to register. Accordingly, we used the state-level telemarketing sales intensity interacted with the lagged cumulative registration rate to capture this "telemarketing diversion" effect.

Finally, "do not call" registration may relate to population heterogeneity – the choices of other people should be more informative if they exhibit similar characteristics (Bikhchandani et al. 1998). Heterogeneity on various dimensions, including age, education, income, race, and religion have been identified as important in social integration (Alesina and Ferrara 2000; McPherson et al. 2001; Costa and Kahn 2003; Alesina et al. 2004; Marmaros and Sacerdote 2006) and new product adoption (Dekimpe et al. 1998; Talukdar et al. 2002; Van den Bulte and Stremersch 2004).

We used these demographic dimensions to construct two new variables that characterized economic and social heterogeneity (see Section 5.2). We further multiplied these two heterogeneity variables by the lagged cumulative registration rate to investigate how population mix and past registrations interacted to affect household decisions. This set of heterogeneity-related measures allowed us to gain a deeper understanding of *why* people followed others' actions, and they, together with the set of "information" and "influence" variables that we described above, comprised the $\bar{z}_{jt}$ vector in (7).

4.2 Data

The key to our contribution was the assembly of a rich data-set, culled from multiple sources.[15] We aggregated the FTC data on "do not call" registrations by county and week to obtain the weekly registrations, r_{jt}. For each county j, we summed the weekly registrations up to week $t-1$, and divided the sum by the estimated number of telephone numbers in a

[15] Please refer to the Appendix for details of the data.

county, M_{j0}, to obtain lagged cumulative registration rates.[16] We computed M_{j0} from Federal Communications Commission (FCC) data on phone number utilization by area code and proportion of residential vis-à-vis non-residential phone lines by state (please refer to the Appendix for details).

From the U.S. Census Bureau and the Association of Religion Data Archives, we obtained county-level demographic data, and we used these data to calculate the heterogeneity-related measures (the detailed procedures are presented in Section 5.2). To characterize social interaction, we needed data on household interaction with other people. We obtained such data from the Social Capital Community Benchmark Survey, which was administered by the Saguaro Seminar at Harvard University.[17] Specifically, we used the responses to questions 51, 56D, 56F, and 56H in the survey as measures of the frequency of interaction with close neighbors, relatives, friends, and co-workers.

To capture the impacts due to different types of learning, we compiled statistics related to exposure to mass media. Specifically, we obtained county-level weekly circulations of newspaper titles in 2003 from the Audit Bureau of Circulation (ABC). We then compiled all newspaper reports of the "do not call" registry from two news archives – Google News and Highbeam Research – from May 2003 onward, and multiplied the numbers of such reports by the county-level circulation of the corresponding newspapers to measure households' exposure to "do not call" reports at a county-week level.[18]

Among the various newspaper reports, we distinguished those that merely mentioned the "do not call" registry from those that also mentioned the *number* of telephone numbers that were already registered (e.g., "people registered more than 10 million phone numbers with the national do-not-call list in its first four days," *San Jose Mercury News*, July 1, 2003). The former variable captures communication of information through an impersonal channel;

---

[16] As explained in Section 4.1, we further multiplied the lagged cumulative registration rate by information, influence, and heterogeneity variables to identify different social effects. As Brock and Durlauf (2007) suggested, using lagged cumulative registrations can help the econometrician identify *endogenous* from *contextual* effects, and hence could reduce the possible bias caused by the reflection problem. For a thorough analysis of the reflection problem and the identification conditions in discrete choice models, see Manski (1993) and Brock and Durlauf (2001, 2007).

[17] For details of the Saguaro Seminar, see http://www.hks.harvard.edu/saguaro/ [Accessed March 17, 2008]. We used the data from the 2000 survey because the 2006 data were not yet available.

[18] Our weighting of newspaper reports by circulation is similar to the concept of gross ratings points (GRP) that is commonly used in marketing (where GRP = reach × frequency). The measures of newspaper reports of the "do not call" registry and the number of people registering were lagged by one week to allow sufficient time for households to act (register) after reading the reports.

the latter variable provides information on the choices of others, and hence could lead to contagion *indirectly* (i.e., it captures indirect social influence through observational learning).

Finally, we compiled consumer telemarketing expenditure at the state level from the Direct Marketing Association. To avoid any potential endogeneity (telemarketers may have adjusted their expenditure after the implementation of the "do not call" registry), we only used telemarketing expenditure data for 2002, and we measured *telemarketing sales intensity* as the *ratio* of the state-level consumer telemarketing expenditure relative to the number of households in a state.

Table 1 reports the descriptive statistics of our data, which spanned the first nine weeks of "do not call" registrations. Table 2 reports their bivariate correlations. Note that the Social Capital Community Benchmark Survey (from which we compiled the frequency of interaction with neighbors, friends, etc.) was administered to only 381 counties which were randomly drawn using proportionate sampling from the total of 3,141 U.S. counties. Accordingly, we restricted our sample to these 381 counties, leading to a balanced panel of 3,429 county-week observations.

[Insert Tables 1 and 2 about here]

5. Results

We first estimated a basic random-effects model that contained only state and time dummy variables. In general, the coefficients of all time dummy variables were significant and decreasing, which is consistent with the observed pattern that "do not call" registrations dropped from the second to the ninth week (see Figure 1). The Breusch and Pagan test of random effects indicated that the random effects specification was preferred ($\chi^2 = 4907$, $p < 0.01$).[19] For brevity, we do not report the detailed results of this specification. All other specifications included state and time dummy variables which we do not report for brevity.

Next, we added the mean household characteristics (the intrinsic preference variables, $\bar{x}_{jt}$) in the estimation. Table 3, specification (i), reports the results. Among the household characteristics, only household size and median household income were significant. Smaller

[19] Both the decreasing "time" effect and the significant Breusch and Pagan test results were consistent across all specifications. We also performed the Hausman's specification test for fixed vis-à-vis random effects, but it produced inconsistent conclusions across model specifications. Nevertheless, in all fixed-effects models, the estimated correlations between the county effects, u_j, and the independent variables were negligible, which was consistent with the key assumption of the random-effects model, that the estimated county heterogeneity was uncorrelated with the explanatory variables.

households and households with higher incomes were more likely to subscribe to the "do not call" registry.

A person in a smaller household is likely to receive, on a per person basis, more telemarketing calls, and also, people with higher income have a higher opportunity cost of time, and hence bear higher costs of receiving telemarketing calls. Therefore, it is reasonable that "do not call" registrations were significantly related to these two characteristics.

[Insert Table 3 about here]

5.1 How information about "do not call" was communicated.

In the next three specifications, we examined *how* communication affected household "do not call" registration. First, we included variables to measure the exposure of households to people who had already subscribed to the "do not call" registry. In particular, we assumed that the impact of communication through social interaction in (7) was represented by

$$\bar{z}'_{jt}\alpha = \sum_{g\in G} \lambda_g I_{jg} + \sum_{g\in G} \alpha_g I_{jg} R_{jt} \ , \qquad (9)$$

where I_{jg} denotes the frequency of households in county j interacting with peer group g, and

$$R_{jt} = \frac{1}{M_{j0}} \sum_{k=1}^{t-1} r_{jk} \qquad (10)$$

represents the lagged cumulative registration rate in county j at time t.

In (9), the product of the social interaction measures with the lagged cumulative registration rate represented the likelihood of interaction with people who had already subscribed to the "do not call" registry. These people could then share information about the "do not call" registry.

Equation (9) also included the social interaction measures as separate explanatory variables. These should be interpreted as being the complements of the products of the social interaction measures with the lagged cumulative registration. They represent the impact of information from people who either had not yet decided whether to subscribe to the "do not call" registry or who had considered and *decided against registering*.

Table 3, specification (ii), reports the results.[20] The coefficients of visits to neighbors and relatives were negative and significant, suggesting that people who frequently visited neighbors and friends who either had not yet decided on "do not call" registry or decided against it were themselves less likely to register. By contrast, the coefficients of visits to neighbors and relatives interacted with the lagged cumulative registration rate were positive and significant, suggesting that people who frequently visited neighbors and friends who had already subscribed to the "do not call" registry were themselves more likely to register.

Overall, the eight social interaction variables made a net positive contribution to registrations, in the sense that the mean predicted utility (the average of the predicted values of the dependent variable using just the eight explanatory variables and their corresponding coefficients) due to these variables was positive.

To more precisely quantify the importance of social interaction in household utility from the "do not call" registry, we compared the variance of their impact against the variance of the impact due to the intrinsic preferences, $\bar{x}_{jt}$. Following Sorensen (2006), the variance due to a set of explanatory variables was calculated as the variance of the predicted value of the dependent variable using just the particular explanatory variables and their corresponding coefficients. To the extent that the variance due to the social interaction variables exceeded the variance due to the intrinsic preference variables, then we infer that social interaction had a stronger influence on household utility from the "do not call" registry.

We report the results of such comparative variance analyses in the last two rows of Table 3, where the standard deviation of social interaction is compared with the standard deviations of intrinsic household preferences, $\bar{x}_{jt}$, and idiosyncratic preferences, v_{jt}. Clearly, the social interaction variables played a stronger role than the intrinsic household preferences, and their effects were almost as large as those of the unobserved idiosyncratic preferences.

Our longitudinal registration data further allowed us to perform a counterfactual experiment – in many studies of social interaction, contemporaneous outcome data from other people in a similar social group were used as one of the key independent variables (e.g., Topa 2001; Conley and Topa 2002; Sorensen 2006). This would generate observations of

[20] We did not include the lagged cumulative registration rate as a separate explanatory variable. Other than social interaction which we already incorporated in (9), and observational learning and telemarketing diversion which we directly identified using other variables (see Section 4.1), there was no good theoretical reason to expect county-level registrations to affect an individual household's registration. Including the lagged cumulative registration rate as a separate variable would distort the coefficients of the other explanatory variables.

"social interaction" which might have been due to omitted variables (see Manski 1993 and Bertrand et al. 2000, and, especially, the discussion of the reflection problem).

To illustrate this potential bias, we re-estimated specification (ii) by adding the current period registration rates, r_{jt}, to the lagged cumulative registration rate in (10). The result was notable. The effects of interacting with close neighbors and relatives who had registered both increased substantially: for neighbors, the coefficient increased from 0.694 to 1.102, while, for relatives, the coefficient increased from 0.027 to 0.045. Further, the overall contribution of the social interaction variables in household utility was much larger: the standard deviation relative to the intrinsic preferences increased from 2.367 to 5.219, while that relative to the idiosyncratic preferences increased from 0.815 to 1.276. Clearly, our use of the lagged cumulative registration rate could mitigate the potential upward bias due to the reflection problem (Brock and Durlauf 2007).

Next, we considered whether information about the "do not call" registry was transmitted through impersonal channels (rather than through social networks such as neighbors, relatives, etc), i.e., information saliency. We included the weekly circulation of all newspapers in the county, and, more importantly, the lagged number of newspaper articles mentioning the "do not call" registry, weighted by circulation, as explanatory variables in the estimation. Table 3, specification (iii), reports the results. Both variables had positive and significant coefficients. However, referring to the comparative variance, newspaper circulation and reports played a less important role than social interaction in household utility from the "do not call" registry.

Newspaper circulation could reflect the effects of other unobserved characteristics, e.g., people who read newspapers may be those who were more concerned about privacy. Hence, circulation may not have precisely measured the impact of mass media. However, the other variable – lagged number of newspaper reports of the "do not call" registry, weighted by circulation – precisely captured the effect of reporting about the registry, and its significance implies that information in the mass media affected household decisions on whether to register.[21]

Finally, in specification (iv), we incorporated both the social interaction and newspaper exposure variables in the estimation. In comparison with specifications (ii) and

[21] In another (unreported) test, we included, from the Social Capital Community Benchmark Survey, the amount of time that people spent in watching TV and browsing the World Wide Web (WWW), but these two variables were insignificant. TV and WWW may not be focused media for information about the "do not call" registry, and so they are noisy measures of information transmission.

(iii), the comparative variance analysis suggested that visits with neighbors, relatives, friends, and co-workers had a bigger impact than newspaper exposure. Apparently, although people did learn about the "do not call" registry from newspapers (and other mass media whose reports were correlated with newspapers), it was social interaction that played a bigger role in communicating information about the "do not call" registry.

5.2 Why contagion in "do not call"?

Having analyzed how information about the "do not call" registry was transmitted, we turned to the issue of why individual households were affected by the registrations of others. We considered two possible motivations for the observed contagion in "do not call" registrations.

The first reason for contagion might have been observational learning – people may have inferred from the registrations of others that the "do not call" registry was beneficial. We compiled the lagged number of newspaper articles that mentioned the *number* of "do not call" registrations, weighted by the corresponding newspaper circulation, and included this as an explanatory variable. Table 3, column (v), reports the results. To control for general newspaper exposure effects, we also included the newspaper circulation and lagged number of "do not call" reports, weighted by circulation.[22]

The overall circulation of newspapers and the lagged number of newspaper articles mentioning the "do not call" registry, weighted by circulation, continued to contribute positively to household utility. However, we did not find evidence of observational learning. In fact, the coefficient of the weighted, lagged number of newspaper articles mentioning the *number* of "do not call" registrations was negative and significant. That is, once people read from newspapers that many people had already registered, their inclination to register actually *decreased*.

Such a negative effect of number of registration is surprising, but it might be explained by consumer *expectations* of improved marketing efficiency. The "do not call" registry raised the operating costs of telemarketers, and so, it might have caused less-desirable telemarketers to exit the market. This would leave the market with a smaller pool

[22] Observational learning could lead to correlations between past and current behavior, which is the underlying assumption in some empirical research, e.g., Santugini (2007). However, as we explained above, other social factors unrelated to learning could also yield such correlation. Hence, instead of using past "do not call" registrations to characterize observational learning, we directly used newspaper reports of the number of registrations. Since we also controlled for exposure to newspapers in general and reports of the "do not call" registry, this variable should precisely identify the effect due to information about the *number* of registrations. Cai et al. (2007) used a similar approach to identify observational learning.

of more desirable telemarketers who are more attractive to consumers (Van Zandt 2004; Anderson and de Palma 2006).

Similarly, people who registered for "do not call" were likely to be the ones who had relatively little interest in buying from telemarketing, and so, their "exit" would actually help telemarketers find the right customers (Hann et al. 2008). These expectations on the supply- and demand-side interactions would have raised household utility from telemarketing, and reduced their utility from the "do not call" registry.

Next, as we explained in Section 4.1, another motivation for households to subscribe to the "do not call" registry was "telemarketing diversion" – as more households registered, telemarketers might divert promotional calls to the remaining households. As such, we specified social effects as

$$\bar{z}'_{jt}\alpha = \alpha T_j R_{jt}, \tag{11}$$

where T_j was the intensity of telemarketing in county j, which was a time-invariant variable at the state level (telemarketing data at the county level were not available).

Table 3, specification (vi), reports the results of this "telemarketing diversion" effect. We did not include telemarketing sales intensity as a separate explanatory variable because it was perfectly collinear with the state dummy variables, and hence its effect was already accounted for in the model. The coefficient of the interaction between telemarketing sales intensity and the lagged cumulative registration was positive and significant. The comparative variance analysis indicated that it contributed significantly to household utility. In fact, its influence was several times larger than that of observational learning about registrations from newspapers.

Taken together, it appears that, although the "do not call" registry might have benefited the telemarketing industry by more accurately matching interested buyers and sellers, on balance, consumers were more eager to sign up and hence avoid promotional calls because of the increased intensity of telemarketing.

Finally, we considered the extent to which household registrations were influenced by other households depended on the closeness of their preferences. The degree of observational learning should be higher among households that are more similar (Bikhchandani et al. 1998).

We calculated the county-level heterogeneity on various demographic dimensions. Following the literature (see, e.g., Alesina and Ferrara 2000), we measured the heterogeneity of race, religion, and education as the probability that any two individuals drawn at random from a county would not belong to the same demographic group. Formally, the heterogeneity on demographic dimension q was

$$H_q = 1 - \sum_{g \in G} s_{qg}^2 , \qquad (12)$$

where s_{qg} is the share of group g in the population on factor q. For age and income inequality, we used the Gini coefficient, which is a standard measure for population dispersion along a metric dimension (see, e.g., Van den Bulte and Stremersch 2004).

Our interest was to study how differences among households affected the informativeness of "do not call" registrations. It was not clear which of the various measures of heterogeneity would be relevant to this issue. In addition, three demographic factors – race, religion, and education – were categorical with probability-based heterogeneity measures, while two – age and income – were cardinal with deviation-based heterogeneity measures (Gini coefficient). We thought it helpful to align these measures.

Accordingly, we conducted a principal component analysis on the five demographic heterogeneity factors to extract their common variances, which then represented intrinsic differences in the population. Table 4 reports the factor loading results. Income, race, and education heterogeneity loaded on one common component, which we called "economic heterogeneity" to reflect their likely correlation with the economic well-being of the people. Religion and age heterogeneity loaded on another component, which we called "social heterogeneity" to reflect their close correspondence to the differences on social dimensions. We created two composite measures – of economic and social heterogeneity – using the predicted factor scores from the principal component analysis. The descriptive statistics of economic and social heterogeneity are reported in Table 1.

Using these two heterogeneity measures, we then checked the extent to which one household's registration was influenced by the registration of other households depended on their economics and social similarity. We specified these effects as,

$$\bar{z}'_{jt} \alpha = \sum_{l} \alpha_l H_l R_{jt} , \qquad (13)$$

where H_l denotes either economic or social heterogeneity. To control for any possible effects due to heterogeneity *per se*, we also included economic and social heterogeneity as separate explanatory variables in the estimation.

Table 3, specification (vii), reports the results. The coefficient of the interaction of social heterogeneity with the lagged cumulative registration rate was significant and *negative*. In communities that were more socially diverse, household "do not call" registration was *less* correlated with the registration of others. This is consistent with the theory that observational learning is decreasing in heterogeneity – the more socially heterogeneous is a community, the less useful are the actions of others as a guide to one's own choice.

It is useful here to contrast our results with those arising from an approach using heterogeneity as a proxy for social interaction under the assumption that social interaction is lower in more heterogeneous communities (e.g., Dekimpe et al. 1998; Talukdar et al. 2002).[23] Using this approach, our results in Table 3, specification (vii), would suggest that "social interaction" had only a slight impact on "do not call" registrations because, by the comparative variance analysis, the influence of economic and social heterogeneity in the utility function – 0.435 relative to intrinsic preferences and 0.249 relative to idiosyncratic preferences – was the lowest among all the variables that we have studied.

By contrast, our results in Table 3, specification (ii), show that the impact of social interaction was more than four times larger in terms of variance relative to intrinsic preferences, and more than three times larger in terms of variance relative to idiosyncratic preferences. Hence, *studies that did not directly observe the extent of social interaction but inferred its significance from proxy variables such as population heterogeneity might draw misleading conclusions on the impact of social interaction*.

Finally, Table 3, specification (viii), included all three sets of variables (observational learning from newspapers, telemarketing diversion, and the impact of heterogeneity) that addressed the issue of why contagion occurs, i.e., why one household's registration was correlated with the registrations of others. Other than some slight differences in statistical significance, the estimation results were mostly consistent with those reported in specifications (v)-(vii). The notable exception was that the coefficient of the interaction of economic heterogeneity with the lagged cumulative registration rate was positive and significant, which seemed to contradict our prior expectation that contagion would be negatively moderated by heterogeneity.

Telemarketing diversion might explain this positive interaction effect. As more households registered, more telemarketing calls would be diverted to the remaining households (Hann et al. 2008). The more heterogeneous are the households in the community, the less well-targeted would be the telemarketing offers, and hence the less attractive would they be to households in the community (Van Zandt 2004; Anderson and de

[23] Referring to Table 2, in our data, economic heterogeneity was somewhat negatively correlated with three of the four measures of social interaction. However, social heterogeneity was somewhat *positively* correlated with the four measures of social interaction.

Palma 2006). Accordingly, in a more (economically) heterogeneous community, registration by others would create a stronger stimulus to registration by the remaining households.[24]

5.3 The "how" and "why" of contagion

Table 3, specification (iv), included measures of social interaction and newspaper reports, which generally explained *how* information about the "do not call" registry was communicated to households. Specification (viii) included measures of observational learning, telemarketing diversion, and the impact of heterogeneity, which generally addressed *why* "do not call" registrations exhibited contagion. For completeness, specification (ix) included all of the variables that we investigated in this study.

Table 5 reports the comparative variance analyses results for each of these five sets of social factors in the combined specification (ix). Previous empirical analyses of contagion variously emphasized social interaction (e.g., Bertrand et al. 2000; Aizer and Currie 2004; Sorensen 2006) and observational learning (e.g., Santugini 2007). In our context, these factors had relatively modest influence in the utility from "do not call" registration. By contrast, telemarketing diversion and the moderating effect of heterogeneity played a larger and more significant role in the utility function.

The immediate implication is that, in any analysis of social factors in individual behavior, it is important to account for social factors specific to the context, as well as the generic social factors – social interaction and observational learning. In our context, the specific social factor was telemarketing diversion.

However, we caution that the relative importance of the various social factors – social interaction, observational learning, and context-specific factors – may vary across contexts. That a specific factor – telemarketing diversion – is relatively more important in the context

[24] Note that telemarketing diversion should also apply to the interaction of social heterogeneity with the lagged cumulative registration rate. Overall, it seems that two forces were in contention – reduced learning owing to heterogeneity (which should decrease registration) and increased telemarketing diversion due to mismatched promotions (which should increase registration). Our results show that the learning effect was stronger for social heterogeneity, whereas the telemarketing diversion effect was stronger for economic heterogeneity. These seem reasonable ex post, because the extent of direct interaction between individuals often vary along social dimensions (Alesina and Ferrara 2000; McPherson et al. 2001; Alesina et al. 2004), and hence the learning effect may dominate in socially heterogeneous communities. Economic heterogeneity, on the other hand, is more "intrinsic" and may not necessarily manifest in direct social interactions. Hence, the negative impact of mismatched promotions may dominate for economically heterogeneous communities, causing people to register. Van den Bulte and Stremersch (2004) also found that income heterogeneity had a positive effect on contagion in new product adoptions.

of the "do not call" registry may have little implication for the relative importance of specific social factors in other settings.

5.4 How much do people value privacy?

The intricate data in this study allowed us to explore how much individuals value privacy from telemarketing. By partitioning the estimated utility, we could examine whether people had a strong intrinsic need for privacy, and whether such need was affected by social interaction, observational learning and telemarketing diversion.

Specifically, from (7) and (8), we can "decompose" the need for privacy as follows:

- The intrinsic need for privacy in terms of household characteristics, $\bar{x}_{jt}$, random effects, u_j, and idiosyncratic preferences, v_{jt};
- Changes in privacy need due to social factors, including social interaction, information saliency, and observational learning;
- Changes in privacy needs due to telemarketing diversion. As mentioned in Section 5.2, because the direct effect of telemarketing was "absorbed" by the state dummy variables, we included the state-specific constants, γ, in computing these changes.

Using the estimates obtained from Table 3, specification (ix), we performed various comparative variance analyses to quantify the relative contributions of the above sets of factors on household demand for privacy. The results are reported in Table 6. Relative to intrinsic need, social interaction, information saliency, and observational learning were rather insignificant. Compared to the standard deviation of the intrinsic needs for privacy, the standard deviation of these social factors was only around 26%.

By contrast, households were much more sensitive to telemarketing diversion. The standard deviation of telemarketing diversion was around 90% of that of the intrinsic need for privacy. Such a finding should not be surprising, because telephone solicitations impose negative externalities on consumers and directly infringe consumer privacy. Overall, it seems that, in the context of telephone marketing, the demand for privacy was more a personal issue that was shaped by intrinsic factors and telemarketing diversion rather than social influences.

Finally, as we stated above, the heterogeneity factors may also capture part of the influence due to telemarketing diversion. Accordingly, we performed another comparative variance analysis by including the set of heterogeneity variables into the effects of telemarketing diversion. The results are reported in the last row of Table 6. Because the heterogeneity variables may have absorbed opposite effects due to telemarketing diversion and learning, on balance, they had little effect on the demand for privacy.

6. Robustness

In the course of producing the estimates reported in Table 3, we made a number of decisions with respect to data and specification. It was important to check the sensitivity of our results to differences in data and specification.

First, we checked whether our results depended on the source of newspaper reports. Instead of using Google News and Highbeam, we separately compiled all newspaper reports of the "do not call" registry and of the number of "do not call" registrations from the Factiva news database. Factiva covered 73 newspapers that were audited by ABC. Table 7, specification (a), reports the results using the Factiva newspaper variables. The coefficients of the newspaper mention variables had the same signs as those in Table 3, specification (ix), and the lagged newspaper reports of number of "do not call" registrations, weighted by circulation was again negative but insignificant.

Next, we tested the robustness with respect to the social interaction variables. Instead of visits to neighbors, relatives, friends, and co-workers, in Table 7, specification (b), we used individual participation in formal and informal social activities, and their interactions with lagged cumulative registrations as measures for social interaction.[25] Both interaction variables were close to significant ($p < 0.10$), but otherwise the signs and significance of all other variables and the overall contribution of social factors to household utility from the "do not call" registry were similar to those in Table 3.

In Table 3, specifications (v), (viii) and (ix), lagged newspaper reports of the number of "do not call" registrations had negative coefficients and mixed significance. Could this result be due a measurement error? Newspaper reports may have had an immediate effect on household registration. By lagging the newspaper measures, the estimated impact on household utility might have been weakened.

To test this, we used contemporaneous newspaper variables in place of the lagged ones. The results are reported in Table 7, specification (c), and they are consistent with those in Table 3. Newspaper reports of the number of "do not call" registrations, weighted by

---

[25] We obtained such data from the Social Capital Community Benchmark Survey. According to the Survey, formal social interaction was defined as the extent of formal group involvements such as attending public meetings and club meetings; informal social interaction was defined as having friends visit home, visiting with relatives, socializing with co-workers outside of work, hanging out with friends in public places, and playing cards and board games.

circulation, continued to be negative and significant. The coefficients of the two newspaper mention variables were somewhat bigger than those reported in Table 3. However, the comparative variance analysis suggests that the contribution of these contemporaneous measures to household utility was marginally smaller than that of the lagged newspaper mention variables.[26]

The limitation of Google News, Highbeam, and Factiva is that they covered only relatively well-circulated newspaper titles. Accordingly, our measures of newspaper reports may have under-stated the impact of mass media in counties dominated by less well-circulated newspaper titles. The latter are probably the more rural counties. To account for this, in Table 7, specification (d), we re-estimated the full specification, excluding those counties that had below median population density.

Most of the social factors had similar signs and significance, albeit their magnitudes differed considerably from those in the earlier specifications because of the trimming of less densely populated counties. More importantly, the lagged newspaper reports of number of "do not call" registrations, weighted by circulation, continued to have a negative but insignificant coefficient. Hence, our results were robust to sample classification based on population density (which more generally reflects county urbanization).

Finally, our dataset included a panel of 381 counties. Household choices in these counties might vary systematically according to local factors which had not been well captured in the demographic variables and random effects that we incorporated in the estimation. It is possible that the model errors were serially correlated. In the next test, we re-estimated the full model by incorporating an AR(1) error structure. The results are reported in Table 7, specification (e). Most of the coefficients had the same signs and similar magnitude, but the model was generally less significant. The relative contribution of the set of social factors to household utility was also lower compared with the other specifications in Tables 3 and 7.

7. Concluding Remarks

Correlation in individual behavior raises two questions – how information is transmitted and why individuals follow the choices of others. We addressed these questions in the context of

[26] We also tried lagging the newspaper mention variables by one more week (i.e., a lag of two weeks rather than one). The signs of the coefficients of the two newspaper mention variables remained the same, but their magnitudes were much smaller as compared with those in Table 3, specification (ix). Together with Table 7, specification (c), this provides strong evidence that the effect of newspaper mentions diminished over time.

the U.S. "do not call" registry, using a rich set of data culled from multiple sources, including longitudinal observations of registration decisions.

Among methods of information transmission, we found that social interaction is relatively more important than news media. Among reasons for contagion, we found that telemarketing diversion is much more important than observational learning, while the extent of learning decreases with social heterogeneity.

We also found that individuals value privacy from telemarketing, and their demand for such privacy was strongly influenced by telemarketing diversion, which imposes negative externalities. By contrast, social interaction, information salience, and observational learning played a relatively minor role in the demand for privacy.

Our results suggest that, in any analysis of social factors in individual behavior, it is important to account for social factors specific to the context, as well as the generic social factors – social interaction and observational learning. However, the relative importance of the various social factors – social interaction, observational learning, and factors specific to the context – may vary across contexts. For example, it is intuitive that the effects of social interaction and observational learning would be stronger for products such as the iPhone, Wikipedia, and movies, and welfare programs.

Previous research did not distinguish between the two effects of social interaction – information and influence. Any discussion of "do not call" with friends or co-workers would surely provide information – the benefits of avoiding telemarketing, how to register, etc. So, social interaction certainly communicates information.

Social interaction might also extend to influence – "you should register … it's so easy", "everyone's registering", etc. In other contexts, such as crime, education, and fashion, social interaction might even extend to peer pressure – "you're not cool if you don't" (see, e.g., Case and Katz 1991).

To account for social interaction, we used measures from the Social Capital Community Benchmark Survey, administered by the Saguaro Seminar. As seemed appropriate, we used these to measure the information conveyed by social interaction. However, we had no way to separately identify the influence effected by social interaction. To this extent, our measures of social interaction provided an upper bound to the impact of social communication of information.

Indeed, any natural experiment is very likely to suffer from the same difficulty of identifying influence in social interaction. If individuals with similar characteristics interact and behave similarly, that would be evidence of social communication of information and, in addition, it could be evidence of peer influence too (Duflo and Saez 2003).

Realistically, it seems difficult to identify the "influence" effect except through a very carefully controlled experiment in which the subjects follow well-defined scripts. All subjects would share information about the item, but only randomly selected participants would exert influence. An alternative is to select a natural setting in which people are already well informed about the item under consideration. Then, any social interaction would likely capture "influence" rather than "information" effect.

The results in Tables 3 and 7 show that most household characteristics were not good predictors of "do not call" registrations, and the idiosyncratic preferences, v_{jt}, contributed most to household utility from the registry. Future research should identify more relevant variables that could explain households' needs for privacy.

Finally, our research characterized observational learning through newspaper reports. This seemed appropriate in the context of "do not call" registration because the government disseminated information on "do not call" mostly through press releases (see, e.g., the FTC's website). However, in other settings, observational learning may occur through different channels and media (Zhang 2006; Cai et al. 2007; Santugini 2007). Future research should evaluate the extent to which newspapers, or, more generally, mass media reports contribute to observational learning.

References

Anderson, Simon P. and André de Palma, "Information Congestion," Working Paper, Dept of Economics, University of Virginia, 2006.

Aizer, Anna and Janet Currie, "Networks or Neighborhoods? Correlations in the Use of Publicly-Funded Maternity Care in California," *Journal of Public Economics*, Vol. 88, No. 12, 2004, 2573-2585.

Alesina, Alberto and Eliana La Ferrara, "Participation in Heterogeneous Communities," *Quarterly Journal of Economics*, Vol. 115, No. 3, 2000, 847-904.

Alesina, Alberto, Reza Baqir and Caroline Hoxby, "Political Jurisdictions in Heterogeneous Communities," *Journal of Political Economy*, Vol. 112, No. 2, April 2004, 348–396.

Beck, Jonathan, "The Sales Effect of Word of Mouth: A Model for Creative Goods and Estimates for Novels," *Journal of Cultural Economics*, Vol. 31, No. 1, 2007, 5-23.

Berry, Steven T., "Estimating discrete-choice models of product differentiation," *RAND Journal of Economics*, Vol. 25, No. 2, Summer 1994, 242-262.

Berry, Steven T., James Levinsohn and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, Vol. 63, No. 4, 1995, 841-890.

Bertrand, Marianne, Erzo F.P. Luttmer and Sendhil Mullainathan, "Network Effects and Welfare Cultures," *Quarterly Journal of Economics*, Vol. 115, No. 3, 2000, 1019-1055.

Bikhchandani, Sushil, David Hirshleifer and Ivo Welch, "A Theory of Fads, Fashions, Custom and Cultural Change as Information Cascades", *Journal of Political Economy*, Vol. 100, 1991, 992-1026.

Bikhchandani, Sushil, David Hirshleifer and Ivo Welch, "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," *Journal of Economic Perspectives*, Vol.12, No. 3, Summer 1998, 151-170.

Brock, William A. and Steven N. Durlauf, "Discrete Choice with Social Interactions," *Review of Economic Studies*, Vol. 68, No. 2, 2001, 235-260.

Brock, William A. and Steven N. Durlauf, "Identification of Binary Choice Models with Social Interactions," *Journal of Econometrics*, Vol. 140, No. 1, 2007, 52-75.

Cai, Hongbin, Yuyu Chen and Hanming Fang, "Observational Learning: Evidence from a Randomized Natural Field Experiment," Guanghua School of Management and IEPR, Peking University, 2008.

Case, Anne and Lawrence Katz, "The Company You Keep: The Effect of Family and Neighborhood on Disadvantaged Youth," NBER Working Paper # 3708, 1991.

Conley, Timothy G. and Giorgio Topa, "Socio-Economic Distance and Spatial Patterns in Unemployment," *Journal of Applied Econometrics*, Vol. 17, 2002, 303-327.

Costa, Dora L. and Matthew E. Kahn, "Civic Engagement and Community Heterogeneity: An Economist's Perspective," *Perspectives on Politics*, Vol. 1, No. 1, March 2003, 103-111.

Dekimpe, Marnik G., Philip M. Parker and Miklos Sarvary, "Staged Estimation of International Diffusion Models: An Application to Global Cellular Telephone Adoption", *Technological Forecasting and Social Change*, Vol. 57, Nos. 1-2, January-February 1998, 105-132.

Duflo, Esther and Emmanuel Saez, "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics*, Vol. 118, No. 3, 2003, 815-842.

Federal Trade Commission, "National Do Not Call registry Opens", News flash, June 27, 2003a. http://www.ftc.gov/opa/2003/06/donotcall.shtm [Accessed April 2, 2008]

Federal Trade Commission, "Thirteen Days Remain for Initial Do Not Call Registration", Press Release, August 19, 2003b. http://www.ftc.gov/opa/2003/08/dnc2weeks.shtm [Accessed April 2, 2008]

Glaeser, Edward L., Bruce I. Sacerdote and Jose A. Scheinkman, "Crime and Social Interactions", *Quarterly Journal of Economics*, Vol. 111, No. 2, May 1996, 507-548.

Goolsbee, Austan and Peter J. Klenow, "Evidence on Learning and Network Externalities in the Diffusion of Home Computers," *Journal of Law and Economics*, Vol. 45, No. 2, 2002, 317-43.

Hann, Il-Horn, Kai-Lung Hui, Yee-Lin Lai, S.Y.T. Lee and I.P.L. Png, "Who Gets Spammed?" *Communications of the ACM*, Vol. 49, No. 10, October 2006, 83-87.

Hann, Il Horn, Kai-lung Hui, Sang-Yong T. Lee and I.P.L. Png, "Consumer Privacy and Marketing Avoidance: A Static Model", *Management Science*, forthcoming, June 2008.

Hann, Il Horn, Kai-Lung Hui, Sang-Yong T. Lee and I.P.L. Png, "Overcoming Online Information Privacy Concerns: An Information Processing Theory Approach," *Journal of Management Information Systems*, Vol. 42, No. 2, 2007, 13-42.

Hui, K.L., "Product Variety under Brand Influence: An Empirical Investigation of Personal Computer Demand," *Management Science*, Vol. 50, No. 5, 2004, 686-700.

Hui, K.L., H.H. Teo and T.S.Y. Lee, "The Value of Privacy Assurance: An Exploratory Field Experiment," *MIS Quarterly*, Vol. 31, No. 1, 2007, 19-33.

Kelly, Morgan and Cormac O. Grada, "Market Contagion: Evidence from the Panics of 1854 and 1857," *American Economic Review*, Vol. 90, No. 5, December 2000, 1110-1124.

Manski, Charles, "Identification of endogenous social effects: the reflection problem", *Review of Economic Studies*, Vol. 60, No. 3, 1993, 531-542.

Marmaros, David and Bruce Sacerdote, "How Do Friendships Form," *Quarterly Journal of Economics*, Vol. 121, No. 1, 2006, 79-119.

McFadden, Daniel, "Modelling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models*. Eds. A. Karlgrist et al. Amsterdam: North-Holland, 1978, 75-96.

McPherson, Miller, Lynn Smith-Lovin and James M. Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, Vol. 27, 2001, 415-444.

Moul, Charles C., "Measuring Word of Mouth's Impact on Theatrical Movie Admissions," *Journal of Economics and Management Strategy*, Vol. 16, No. 4, 2007, 859-892.

Plumper, Thomas and Vera E. Troeger, "Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects," *Political Analysis*, Vol. 15, 2007, 124-139.

Sacerdote, Bruce I., "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics*, Vol. 116, No. 2, 2001, 681–704.

Santugini, Marc, "Observational Learning in the Motion Picture Market", Department of Economics, University of Virginia, January 11, 2007.

Soetevent, Adriaan R., "Empirics of the Identification of Social Interactions: An Evaluation of the Approaches and their Results," *Journal of Economic Surveys*, Vol. 20, No. 2, 2006, 193-236.

Sorensen, Alan T., "Social Learning and Health Plan Choice," *RAND Journal of Economics*, Vol. 37, No. 4, 2006, 929-945.

Talukdar, Debabrata, K. Sudhir and Andrew Ainslie, "Investigating New Product Diffusion Across Products and Countries," *Marketing Science*, Vol. 21, No. 1, 2002, 97-114.

Topa, Giorgio, "Social Interactions, Local Spillovers and Unemployment," *Review of Economic Studies*, Vol. 68, No. 2, 2001, 261-295.

Van den Bulte, Christophe and Gary L. Lilien, "Medical Innovations Revisited: Social Contagion versus Marketing Effort," *American Journal of Sociology*, Vol. 106, No. 5, 2001, 1409-1435.

Van den Bulte, Christophe and Stefan Stremersch, "Social Contagion and Income Heterogeneity in New Product Diffusion: A Meta-analytic Test," *Marketing Science*, Vol. 23, No. 4, 2004, 530–544.

Van Zandt, Timothy, "Information Overload in a Network of Targeted Communication," *RAND Journal of Economics*, Vol. 35, No. 3, Autumn 2004, 542-560.

Varian, Hal, Fredrik Wallenberg and Glenn Woroch, "The Demographics of the Do-Not-Call List", *IEEE Security & Privacy*, Vol. 3, No. 1, January/February 2005, 34-39.

Varian, Hal, Fredrik Wallenberg and Glenn Woroch, "Who Signed Up for the Do-Not-Call List?" University of California, Berkeley, June 2004.

Zhang, Juanjuan, "The Sound of Silence: Evidence of Observational Learning from the U.S. Kidney Market," MIT Sloan School of Management, July 8, 2006.

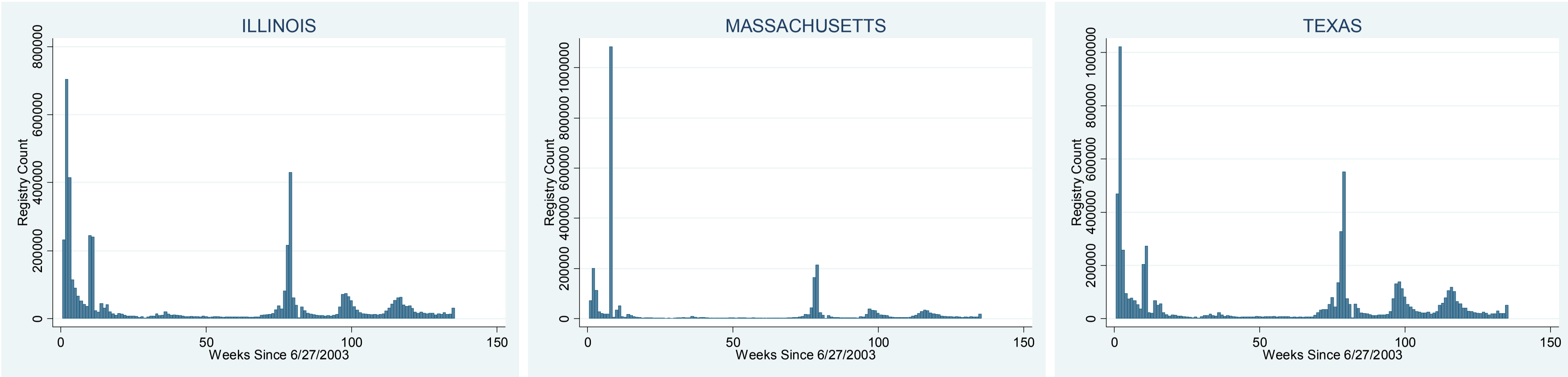

Figure 1. Registrations of fixed-line telephone numbers from (a) Illinois; (b) Massachusetts; and (c) Texas

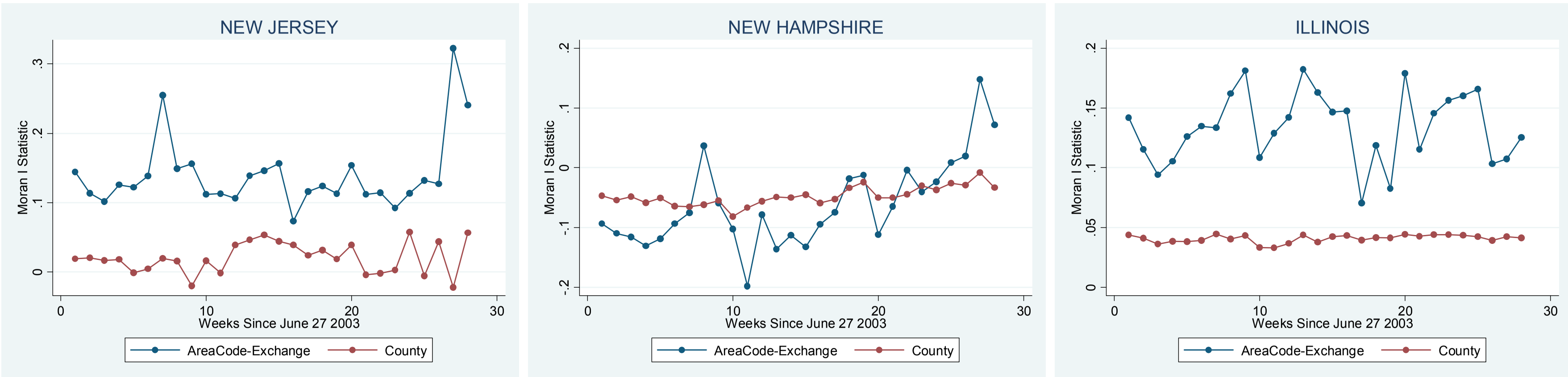

Figure 2. Spatial autocorrelations in (a) New Jersey; (b) New Hampshire; and (c) Illinois

Table 1. Descriptive Statistics

	Variable	Source	Mean	Std. Dev.	Min	Max
1	Weekly registration	FTC	2,901.97	9,762.88	2.00	239,772.00
2	Initial market potential (M_{j0}; in thousands)[a]	calculated	204.70	477.11	1.70	5,749.47
3	Lagged cumulative registrations (as a proportion of total market potential)	calculated	0.09	0.06	0	0.39
4	Average number of people in a household	Census	2.64	0.20	2.07	3.69
5	Median household income (in thousand $)	Census	40.95	10.13	21.52	81.05
6	Percentage of people who were economically active	Census	0.77	0.02	0.70	0.84
7	Percentage of people above 16 years old who were employed	Census	0.97	0.01	0.91	0.99
8	Percentage of people below 65 years old	Census	0.87	0.03	0.74	1.00
9	Percentage of households with female head	Census	0.11	0.04	0.04	0.28
10	Percentage of people above 25 years old who received at least high school education	Census	0.81	0.07	0.52	0.94
11	Percentage of people who were linguistically isolated	Census	0.02	0.02	0	0.17
12	Percentage of households who owned the housing units in which they lived	Census	0.72	0.09	0.31	0.88
13	Frequency of visiting immediate neighbors[b]	Saguaro	5.17	1.21	1.00	7.00
14	Frequency of visiting relatives[b]	Saguaro	27.87	16.76	0	60.00
15	Frequency of having friends visiting[b]	Saguaro	22.20	15.60	0	60.00
16	Frequency of socializing with co-workers outside work[b]	Saguaro	14.09	14.39	0	60.00
17	Weekly newspaper circulation (in thousands)	ABC	0.31	0.11	0.02	0.67
18	Newspaper reports of DNC (weighted by weekly circulation; in thousands)[c]	Google	0.58	1.46	0	17.01
19	Newspaper reports of *number* of DNC registrations (weighted by weekly circulation; in thousands)[c]	Google	0.13	0.50	0	10.93
20	Telemarketing sales intensity (in thousand US$)[d]	DMA	1.70	0.23	1.30	3.34
21	Economic heterogeneity[e]	calculated	0.24	0.92	-2.31	2.50
22	Social heterogeneity[e]	calculated	0.14	0.84	-3.47	2.44

[a] Please refer to Appendix for formula.

[b] From the Social Capital Community Benchmark Survey – see http://www.hks.harvard.edu/saguaro/communitysurvey/index.html.

[c] For each county, we multiplied the number of newspaper articles that reported the “do not call” registry (or the number of telephone numbers on the registry) by the corresponding weekly circulation.

[d] Computed by dividing the 2002 state-level telemarketing expenditure (in thousand US$) by the total number of households in the state.

[e] We computed economic heterogeneity by using the factor scores obtained from a principal-component analysis of income, race, and education heterogeneity. We computed social heterogeneity by using the factor scores obtained from a principal-component analysis of religion and age heterogeneity.

Table 2. Correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1.00																					
2	0.61	1.00																				
3	-0.18	-0.04	1.00																			
4	0.00	0.02	-0.08	1.00																		
5	0.19	0.28	0.10	0.27	1.00																	
6	-0.10	-0.15	-0.15	0.31	-0.47	1.00																
7	0.00	-0.03	0.15	-0.13	0.47	-0.31	1.00															
8	0.12	0.18	0.02	0.58	0.46	-0.27	-0.03	1.00														
9	0.08	0.14	-0.08	0.08	-0.30	0.19	-0.53	0.17	1.00													
10	0.09	0.12	0.15	-0.04	0.59	-0.59	0.39	0.23	-0.53	1.00												
11	0.18	0.32	-0.12	0.36	0.12	0.13	-0.30	0.24	0.18	-0.21	1.00											
12	-0.16	-0.27	0.03	0.14	0.12	0.30	0.36	-0.23	-0.42	-0.03	-0.34	1.00										
13	-0.03	-0.04	0.04	0.02	-0.03	-0.01	0.12	-0.02	-0.14	0.15	-0.07	0.08	1.00									
14	-0.03	-0.04	-0.02	0.12	-0.01	0.10	0.01	0.01	0.00	-0.06	0.01	0.10	0.16	1.00								
15	0.00	0.01	0.03	0.06	0.01	-0.05	-0.02	0.03	-0.11	0.14	0.01	0.01	0.21	0.29	1.00							
16	-0.02	-0.03	0.00	0.04	-0.05	0.01	-0.11	0.00	-0.02	0.02	0.05	-0.03	0.14	0.16	0.28	1.00						
17	0.13	0.19	0.14	0.06	0.40	-0.22	0.12	0.13	-0.07	0.33	0.07	-0.13	0.05	0.01	0.01	0.01	1.00					
18	0.18	0.14	-0.16	0.01	0.22	-0.17	0.07	0.17	0.09	0.05	0.03	-0.09	-0.02	-0.05	-0.01	-0.02	0.10	1.00				
19	0.08	0.07	-0.23	0.04	0.10	-0.06	0.03	0.09	0.04	0.02	0.01	-0.02	0.01	-0.02	0.01	0.01	0.06	0.66	1.00			
20	0.05	0.09	0.06	0.08	0.23	-0.20	0.11	0.02	-0.13	0.21	-0.09	0.00	0.05	0.04	0.08	0.03	0.16	0.08	0.03	1.00		
21	0.16	0.26	-0.09	0.10	-0.15	0.13	-0.48	0.23	0.69	-0.36	0.47	-0.55	-0.10	-0.05	-0.06	0.01	-0.03	0.07	0.04	-0.17	1.00	
22	0.08	0.10	0.01	0.46	0.19	-0.07	-0.14	0.38	0.19	0.09	0.32	-0.12	0.03	0.09	0.01	0.04	0.19	0.07	0.04	0.12	0.10	1.00

Note: Variable names are reported in Table 1.

Table 3. Random-effects estimation

	(i)	(ii)	(iii)	(iv)
Frequency of visiting immediate neighbors		-0.047*** (0.011)		-0.049*** (0.011)
Frequency of visiting relatives		-0.003*** (0.001)		-0.003*** (0.001)
Frequency of having friends visiting		0.002 (0.001)		0.002 (0.001)
Frequency of socializing with co-workers outside work		-0.001 (0.001)		-0.001 (0.001)
Freq. of visiting immediate neighbors × lagged cumulative registration rate		0.694*** (0.060)		0.669*** (0.060)
Freq. of visiting relatives × lagged cumulative registration rate		0.027*** (0.008)		0.026*** (0.008)
Freq. of having friends visiting × lagged cumulative registration rate		-0.007 (0.009)		-0.009 (0.009)
Freq. of socializing with co-workers × lagged cumulative registration rate		0.003 (0.009)		0.003 (0.009)
Weekly newspaper circulation			0.796*** (0.213)	0.537*** (0.112)
Lagged newspaper reports of DNC, weighted by circulation			0.010** (0.005)	0.012** (0.005)
Lagged newspaper reports of *number* of DNC registrations, weighted by circulation				
Telemarketing sales intensity × lagged cumulative registration rate				
Economic heterogeneity				
Social heterogeneity				
Economic heterogeneity × lagged cumulative registration rate				
Social heterogeneity × lagged cumulative registration rate				
Average number of people in a household	-0.426** (0.215)	-0.275** (0.107)	-0.444** (0.211)	-0.290*** (0.106)
Median household income	0.011** (0.006)	0.008*** (0.003)	0.006 (0.006)	0.004 (0.003)
Percentage of people who were economically active	0.088 (1.930)	-0.273 (0.950)	0.225 (1.901)	-0.096 (0.948)
Percentage of people above 16 years old who were employed	-1.689 (3.215)	-1.187 (1.621)	-1.680 (3.161)	-1.089 (1.612)
Percentage of people below 65 years old	0.011 (1.376)	-0.060 (0.672)	0.756 (1.365)	0.461 (0.676)
Percentage of households led by female	0.407 (0.996)	0.160 (0.509)	0.116 (0.981)	-0.025 (0.505)
Percentage of people above 25 years old who received at least high school education	0.654 (0.778)	0.125 (0.378)	0.546 (0.767)	0.100 (0.377)
Percentage of people who were linguistically isolated	-3.331 (1.708)	-2.933*** (0.827)	-2.969 (1.689)	-2.654*** (0.830)
Percentage of owner-occupier households	0.122 (0.491)	-0.014 (0.246)	0.433 (0.489)	0.214 (0.249)
Overall R^2	0.777	0.837	0.7823	0.838

$s.d.(\bar{z}'_{jt}\alpha)/s.d.(\bar{x}'_{jt}\beta)$	n.a.	2.367	0.653	2.810
$s.d.(\bar{z}'_{jt}\alpha)/s.d.(v_{jt})$	n.a.	0.815	0.302	0.832

*** $p < 0.01$; ** $p < 0.05$; robust standard errors are reported in parentheses.

Notes:

1. Dependent variable: $\log(r_{jt}) - \log\left(M_{j0} - \sum_{k=1}^{t} r_{jk}\right)$
2. All specifications included state and year dummy variables.

Table 3 – continued

	(v)	(vi)	(vii)	(viii)	(ix)
Frequency of visiting immediate neighbors					-0.010 (0.012)
Frequency of visiting relatives					-0.003** (0.001)
Frequency of having friends visiting					0.003** (0.001)
Frequency of socializing with co-workers outside work					-0.001 (0.001)
Freq. of visiting immediate neighbors × lagged cumulative registration rate					0.179 (0.095)
Freq. of visiting relatives × lagged cumulative registration rate					0.023*** (0.008)
Freq. of having friends visiting × lagged cumulative registration rate					-0.019** (0.009)
Freq. of socializing with co-workers × lagged cumulative registration rate					0.005 (0.009)
Weekly newspaper circulation	0.798*** (0.214)			0.462*** (0.106)	0.452*** (0.104)
Lagged newspaper reports of DNC, weighted by circulation	0.021*** (0.005)			0.022*** (0.005)	0.020*** (0.005)
Lagged newspaper reports of *number* of DNC registrations, weighted by circulation	-0.041** (0.017)			-0.032 (0.020)	-0.030 (0.020)
Telemarketing sales intensity × lagged cumulative registration rate		2.935*** (0.144)		3.153*** (0.143)	2.579*** (0.327)
Economic heterogeneity			0.003 (0.052)	-0.056 (0.029)	-0.056 (0.029)
Social heterogeneity			0.024 (0.039)	0.022 (0.024)	0.027 (0.024)
Economic heterogeneity × lagged cumulative registration rate			0.138 (0.135)	0.651*** (0.155)	0.686*** (0.159)
Social heterogeneity × lagged cumulative registration rate			-1.023*** (0.138)	-1.101*** (0.155)	-1.096*** (0.155)
Average number of people in a household	-0.440** (0.212)	-0.294*** (0.102)	-0.312 (0.218)	-0.160 (0.101)	-0.144 (0.099)
Median household income	0.006 (0.006)	0.007** (0.003)	0.012** (0.006)	0.004 (0.003)	0.004 (0.003)
Percentage of people who were economically active	0.260 (1.902)	-0.131 (0.907)	-0.624 (1.958)	-0.707 (0.893)	-0.726 (0.881)
Percentage of people above 16 years old who were employed	-1.667 (3.165)	-0.569 (1.535)	-1.238 (3.247)	-0.409 (1.526)	-0.563 (1.528)
Percentage of people below 65 years old	0.753 (1.367)	0.130 (0.643)	-0.274 (1.368)	0.279 (0.635)	0.210 (0.617)
Percentage of households led by female	0.122 (0.981)	0.135 (0.483)	0.819 (1.179)	0.618 (0.556)	0.502 (0.542)
Percentage of people above 25 years old who received at least high school education	0.553 (0.767)	0.165 (0.356)	0.700 (0.770)	0.138 (0.345)	0.086 (0.340)
Percentage of people who were linguistically isolated	-2.988 (1.691)	-2.804*** (0.801)	-3.365 (1.728)	-2.369*** (0.816)	-2.352*** (0.790)
Percentage of owner-occupier households	0.436 (0.490)	-0.007 (0.237)	0.186 (0.524)	0.222 (0.246)	0.213 (0.241)
Overall R^2	0.782	0.842	0.777	0.849	0.852

$s.d.(\bar{z}'_{jt}\alpha) / s.d.(\bar{x}'_{jt}\beta)$	0.671	2.896	0.435	4.109	4.465
$s.d.(\bar{z}'_{jt}\alpha) / s.d.(v_{jt})$	0.309	0.954	0.249	1.054	1.085

*** $p < 0.01$; ** $p < 0.05$; robust standard errors are reported in parentheses.

Notes:

1. Dependent variable: $\log(r_{jt}) - \log\left(M_{j0} - \sum_{k=1}^{t} r_{jk}\right)$
2. All specifications included state and year dummy variables.

Table 4. Factor loadings

Variable	Factor 1	Factor 2
Income heterogeneity	0.699	
Race heterogeneity	0.847	
Education heterogeneity	0.864	
Religion heterogeneity		0.580
Age heterogeneity		0.810

Note: Only loadings with absolute value > 0.4 are reported.

Table 5. Relative contributions in utility — specification (ix)

Factors	$s.d.(\bar{z}'_{jt}\alpha)/s.d.(\bar{x}'_{jt}\beta)$	$s.d.(\bar{z}'_{jt}\alpha)/s.d.(v_{jt})$
Social interaction	1.005	0.244
Newspaper reports	0.741	0.180
Observational learning	0.183	0.044
Telemarketing diversion	3.469	0.843
Impact of heterogeneity	1.035	0.251

Table 6. Relative need for privacy

Comparisons	Comparative variance
Social factors (including social interaction, information saliency, and observational learning) relative to intrinsic need for privacy	0.258
Changes in privacy value due to telemarketing diversion relative to intrinsic need for privacy	0.901
Changes in privacy value due to telemarketing diversion (including the effects of heterogeneity) relative to intrinsic need for privacy	0.891

Note: the intrinsic need was measured by $\bar{x}_{jt}$, u_j, and v_{jt}.

Table 7. Robustness

	(a)	(b)	(c)	(d)	(e)
Frequency of visiting immediate neighbors	-0.015 (0.012)		-0.010 (0.012)	-0.015 (0.012)	-0.001 (0.014)
Frequency of visiting relatives	-0.003** (0.001)		-0.003** (0.001)	-0.004*** (0.001)	-0.003** (0.001)
Frequency of having friends visiting	0.003** (0.001)		0.002** (0.001)	0.002 (0.001)	0.003** (0.001)
Frequency of socializing with co-workers outside work	-0.001 (0.001)		-0.001 (0.001)	0.001 (0.001)	-0.001 (0.001)
Formal social interaction		-0.015 (0.020)			
Informal social interaction		0.054 (0.034)			
Freq. of visiting immediate neighbors × lagged cumulative registration rate	0.210** (0.097)		0.183 (0.096)	0.098 (0.095)	0.084 (0.114)
Freq. of visiting relatives × lagged cumulative registration rate	0.023*** (0.008)		0.022*** (0.008)	0.039*** (0.009)	0.017** (0.009)
Freq. of having friends visiting × lagged cumulative registration rate	-0.019** (0.009)		-0.018** (0.009)	-0.017** (0.009)	-0.022** (0.009)
Freq. of socializing with co-workers × lagged cumulative registration rate	0.005 (0.009)		0.005 (0.009)	-0.002 (0.009)	0.007 (0.009)
Formal social interaction × lagged cumulative registration rate		0.278 (0.152)			
Informal social interaction × lagged cumulative registration rate		-0.516 (0.288)			
Weekly newspaper circulation	0.420*** (0.107)	0.447*** (0.107)	0.451*** (0.104)	0.154 (0.084)	0.539*** (0.131)
Lagged newspaper reports of DNC, weighted by circulation		0.023*** (0.005)		0.020*** (0.005)	0.010 (0.007)
Lagged newspaper reports of *number* of DNC registrations, weighted by circulation		-0.033 (0.020)		-0.025 (0.022)	-0.020 (0.016)
Newspaper reports of DNC, weighted by circulation			0.027*** (0.005)		
Newspaper reports of *number* of DNC registrations, weighted by circulation			-0.053*** (0.018)		
Lagged newspaper reports of DNC, weighted by circulation (Factiva)	0.090** (0.045)				
Lagged newspaper reports of *number* of DNC registrations, weighted by circulation (Factiva)	-0.192 (0.101)				
Telemarketing sales intensity × lagged cumulative registration rate	2.471*** (0.332)	3.223*** (0.142)	2.588*** (0.328)	3.457*** (0.323)	2.341*** (0.374)
Economic heterogeneity	-0.056 (0.030)	-0.059** (0.029)	-0.057** (0.029)	-0.047 (0.027)	-0.013 (0.034)
Social heterogeneity	0.039 (0.024)	0.022 (0.024)	0.025 (0.024)	0.069*** (0.024)	0.025 (0.026)
Economic heterogeneity × lagged cumulative registration rate	0.668*** (0.163)	0.653*** (0.159)	0.698*** (0.160)	0.920*** (0.171)	0.277 (0.178)
Social heterogeneity × lagged cumulative registration rate	-1.147*** (0.155)	-1.116*** (0.155)	-1.077*** (0.155)	-0.910*** (0.169)	-1.159*** (0.178)
Average number of people in a household	-0.144	-0.149	-0.142	0.004	-0.177

	(0.101)	(0.101)	(0.099)	(0.092)	(0.129)
Median household income	0.004 (0.003)	0.005 (0.003)	0.004 (0.003)	0.000 (0.002)	0.005 (0.003)
Percentage of people who were economically active	-1.090 (0.902)	-0.650 (0.881)	-0.744 (0.884)	0.343 (0.771)	-0.523 (1.090)
Percentage of people above 16 years old who were employed	-0.300 (1.553)	-0.414 (1.513)	-0.565 (1.522)	-0.764 (1.248)	-0.441 (1.814)
Percentage of people below 65 years old	0.056 (0.631)	0.120 (0.631)	0.204 (0.617)	-0.578 (0.484)	0.195 (0.774)
Percentage of households led by female	0.665 (0.550)	0.656 (0.552)	0.513 (0.541)	-0.755 (0.446)	0.549 (0.695)
Percentage of people above 25 years old who received at least high school education	0.037 (0.342)	0.103 (0.343)	0.084 (0.339)	0.082 (0.284)	0.213 (0.414)
Percentage of people who were linguistically isolated	-2.409*** (0.801)	-2.511*** (0.813)	-2.352*** (0.788)	-2.195*** (0.637)	-2.666*** (0.999)
Percentage of owner-occupier households	0.249 (0.249)	0.177 (0.244)	0.212 (0.240)	0.134 (0.230)	0.175 (0.302)
Overall R^2	0.852	0.850	0.852	0.881	0.843
$s.d.(\bar{z}'_{jt}\alpha)/s.d.(\bar{x}'_{jt}\beta)$	4.159	4.107	4.451	6.109	3.062
$s.d.(\bar{z}'_{jt}\alpha)/s.d.(v_{jt})$	1.074	1.073	1.083	1.341	0.984

*** $p < 0.01$; ** $p < 0.05$; robust standard errors are reported in parentheses.

Notes:

1. Dependent variable: $\log(r_{jt}) - \log\left(M_{j0} - \sum_{k=1}^{t} r_{jk}\right)$
2. All specifications included state and year dummy variables.
3. For specification (d), sample size = 2,790.
4. An AR(1) error structure was specified for specification (e).

Appendix. Data and Variable Construction

Weekly "do not call" registrations

We obtained records of registrations with the federal "do not call" registry from the FTC for the period between June 27, 2003 and January 6, 2006. These records showed registrations by redacted telephone number for each area code and exchange, e.g., (617) 363-xxxx, by date of registration. For each area code and exchange, we aggregated the FTC daily-level data to the weekly level to obtain the number of "do not call" registrations for each calendar week, defined as Sunday to Saturday.[1]

To proceed, we needed to match the registrations with data on demographics, social interaction, and observational learning which were available at the county level. We procured the *North American Local Exchange NPA-NXX Database* (NALENND) from Quentin Sager Consulting[2]. Using the NALENND database, we identified the counties served by each telephone exchange.

Some telephone exchanges spanned multiple counties. For these exchanges, we used the NALENND database to allocate the "do not call" registrations within an exchange to the respective counties according to the relative number of households in the counties as reported by Census 2000. Additionally, we tried two other methods of allocating the exchange-level registrations to the respective counties: number of housing units and the population in each county. Across these three methods of allocation, our findings from the empirical analyses were remarkably similar.

Based on information from the NALENND database, we also removed "do not call" registrations originating from area codes and exchanges associated with mobile phones, pagers, and federal, state and local government.

Market potential of registrations

There are no published measures or statistics associated with the county-level initial market potential (M_{j0}) for "do not call" registrations. As such, we had to rely on related

---

[1] Since "do not call" registrations started on June 27, 2003 which was a Friday, registrations for the first week in our data set comprised registrations from just two days -- June 27 and 28.
[2] http://www.quentinsagerconsulting.com/npanxx_phonecodes.htm [Accessed April 8, 2008]

measures published by the Federal Communications Commission (FCC) in order impute a reasonably accurate measure of the initial market potential.

Specifically, we used data on the percentages of phone numbers in use by area code (Stroup and Vu (2003), Table 6)[3] and the percentages of phone numbers associated with residential, rather than non-residential, subscribers by state (FCC (2003), Table 11)[4]. On the assumption that these percentages were uniform across all exchanges within area codes and states respectively, we then computed the initial market potential for each area code exchange-county as (10,000 lines × relative number of households in the county[5] × percentage of numbers in use × percentage of residential numbers). The county-level initial market potential, M_{j0}, was then derived by aggregating across all exchanges within a county.

<u>Weekly newspaper circulation</u>

We computed the county-level weekly circulation of newspaper titles in the fall of 2003 based on circulation as reported by the Audit Bureau of Circulation (ABC). For each newspaper title, the ABC reported the "coverage", in terms of the ratio of circulation to the number of households, for possibly four days of the week – Monday, Friday, Saturday, and Sunday – by county.

Then, for each newspaper, according to its respective publication cycle, we computed the *weekly* circulation, by county, as follows:

$$\text{Weekly circulation} = \begin{cases} \text{Monday} \times 5, \text{ if only Monday circulation reported} \\ \text{Monday} \times 4 \ + \ \text{Friday, if only Monday and Friday circulation reported} \\ \text{Monday} \times 4 \ + \ \text{Friday} \ + \ \text{Saturday, if Sunday circulation not reported} \\ \text{Monday} \times 4 \ + \ \text{Friday} \ + \ \text{Saturday} \ + \ \text{Sunday, if all days' circulation reported} \end{cases}$$

We summed the weekly circulation across all newspapers with circulation in a county to derive the weekly circulations of all newspapers for that county.

---

[3] http://www.fcc.gov/Bureaus/Common_Carrier/Reports/FCC-State_Link/IAD/utilizationjun2003.pdf [Accessed April 8, 2008]

[4] http://www.fcc.gov/Bureaus/Common_Carrier/Reports/FCC-State_Link/IAD/lcom1203.pdf [Accessed April 8, 2008]

[5] Recall that some exchanges span multiple counties. Thus, we had to apportion the maximum number of 10,000 phone lines within an exchange across the various counties, according to the relative number of households in the county.

Newspaper reports of "do not call" registry and registration

We compiled newspaper reports of the "do not call" registry from three proprietary news archives – Factiva, Google News, and Highbeam Research. The combination of these three resources covered over 1200 online news sources. However, among these 1200 news sources, only 170 from Google News and 73 from Factiva were audited by the ABC.

Our primary news data set was compiled from Google News and Highbeam Research. The data retrieval process began from a search on "do not call", following which a web crawler parsed the webpage search results into XML data formats. The objective of this retrieval step was to gather as many relevant documents as possible. It retrieved around 400,000 reports. Next, these reports were progressively filtered by successively adding constraint words such as "do not call" with "subscribers", and "do not call" with "millions". The filtering process narrowed the set of news reports to approximately 20,000.

For each of the news sources, we then counted the number of newspaper reports in each calendar week covering news of the "do not call" registry and the number of telephone lines registered with the registry. We then retained only observations of news sources which were audited by ABC. For these ABC-audited news sources, we then aggregated the weekly counts of newspaper reports to the county level by summing the counts across all newspapers distributed in a county.

In order to derive a county-specific news exposure measure similar to the gross ratings point measure used in marketing research, we then multiplied the weekly counts of "do not call" registry reports for each newspaper with the corresponding weekly circulation of the newspaper in the county.

We followed similar steps using Factiva. As Factiva covered fewer ABC-audited newspapers, the numbers of news reports were lower than those based on Google News and Highbeam Research.